

# ホワイトペーパー

## Fujitsu PRIMERGY サーバ

### 512E HDD : テクノロジー、パフォーマンス、構成

物理 4K セクターを使用した HDD の新技術により、ストレージ密度、パフォーマンス、およびエラー訂正機能が向上されています。その新技術の HDD が現在、PRIMERGY サーバに搭載されています。これは、外部に対して 512 バイトの論理セクターをエミュレートした形にもなっています。最適なパフォーマンスを得るには、関連するソフトウェアコンポーネントに適切なバージョンを使用する必要があります。

#### バージョン

1.2a  
2014-01-29



## 目次

ドキュメントの履歴 .....	2
テクノロジーと使用例 .....	3
パフォーマンス .....	4
高パフォーマンスサポートを実現するための 512e ハードディスクの推奨構成 .....	6
Windows オペレーティングシステムのインストール .....	8
関連資料 .....	9
お問い合わせ先 .....	9

## ドキュメントの履歴

### バージョン 1.0 (2013-06-06)

初版

### バージョン 1.0a (2013-06-12)

マイナー修正

### バージョン 1.0b (2013-06-18)

ソフトウェアコンポーネントのバージョンを修正

### バージョン 1.1 (2013-12-18)

オペレーティングシステムのインストールに関する新しい章  
マイナー修正

### バージョン 1.2 (2014-01-28)

SVIM のバージョンを修正

### バージョン 1.2a (2014-01-29)

SVIM のバージョンを修正

## テクノロジーと使用例

長期にわたり、HDD の物理セクターの構成は、各セクターに 512 バイトのユーザーデータと一定のオーバーヘッド（「エラー訂正コード」など）を含むものでした。しかし、このセクターフォーマットは近年、その限界に達しています。今日要求されるストレージ密度、パフォーマンス、およびエラー訂正機能の条件に対応できなくなっています。これらの限界を克服するため、2009 年に各社共通の形式となる 4096 バイトの物理セクターフォーマットが定められました（「アドバンスドフォーマット」、4K セクター）。このフォーマットを使用するには、影響を受けるあらゆるハードウェアおよびソフトウェアコンポーネントをサポートすることが前提とされます。これらのコンポーネントには次のものがあります。

HDD、ディスクコントローラー、オペレーティングシステム、ドライバ、アプリケーション、その他ツール。

メーカーによるこれらすべてのコンポーネントの移行には当然ながらある程度時間がかかるため、暫定的に 512 バイトセクターのエミュレーションテクノロジーが導入されています。つまり、HDD の内部は 4K 物理セクターでありながら、外部から見ると論理 512 バイトセクターをエミュレートしたものとして扱われます。これにより、新しいハードディスクも既存の環境で使用することが可能になります。これらの HDD は、短い名前「512e」で呼ばれることもあります（「e」は「エミュレーション」を意味する）。

現在 PRIMERGY サーバに搭載されるハードディスクでは、4K テクノロジーは 512e HDD の形でのみ実現されています。この場合、512 バイトセクターのエミュレーションは常に有効化され、4K ネイティブへのスイッチモードは存在しません。

現在、次の 512e HDD が PRIMERGY サーバ用に提供されています（SATA のみの対応）。

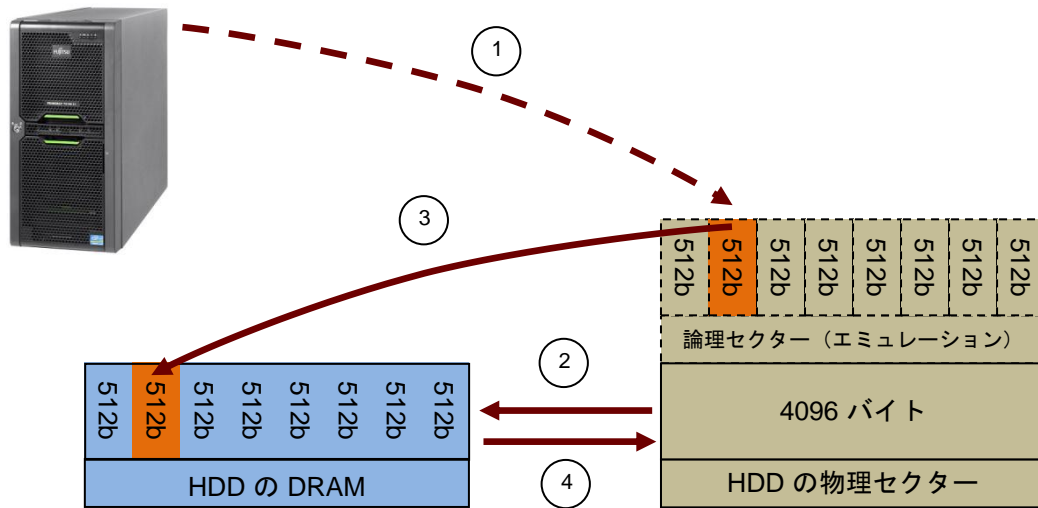
発注コード	説明	出荷日
S26361-F3797-E100/L100 *)	HD SATA 6G 1TB 5.4K HOT PL 2.5" ECO	2013 年 6 月 17 日
S26361-F3798-E100/L100	HD SATA 6G 1TB 5.4K HOT PL 2.5" ECO	2013 年 6 月 17 日
S26361-F3700-E250/L250	HD SATA 6G 250GB 7.2K HOT PL 3.5" ECO	2012 年 8 月
S26361-F3700-E500/L500	HD SATA 6G 500GB 7.2K HOT PL 3.5" ECO	2012 年 8 月
S26361-F3701-E250/L250	HD SATA 6G 250GB 7.2K HOT PL 3.5" ECO	2012 年 8 月
S26361-F3701-E500/L500	HD SATA 6G 500GB 7.2K HOT PL 3.5" ECO	2012 年 8 月
S26361-F3574-E250/L250 **)	HD SATA 3G 250GB 7.2K NO HOT PL 3.5"	2012 年 8 月
S26361-F3293-E250/L250	HD SATA 3G 250GB 7.2K HOT PL 3.5" ECO	2012 年 8 月

\*) TX140 S1 (S26361-F3554-E8、RAID Ctrl SAS 6G 0/1 (D2607) ) 用の特別リリース

\*\* ) S26361-F3574-E250 – 販売終息

国または販売地域によっては、一部のコンポーネントが利用できない場合があります。

アクセスは物理 4K セクターの一部のみの処理となり、現在これらの HDD でのエミュレーションにより正しく機能します。パフォーマンスへの影響は、4K セクター部分への書き込み処理にのみ発生します。詳しく見ると、物理セクターの「エラー訂正コード」の書き換えはセクターの 4K ユーザーデータ全体に対してのみ可能であることから、この場合に HDD はあらかじめ 4K セクターの内容全体を読み取る必要があります。つまり、「読み取り-訂正-書き込み」のサイクルを実行する必要があります。



- 1) ホストからの 512 バイト (「512b」) セクターの書き込み要求
- 2) HDD が 4KB セクターを DRAM に読み込み
- 3) HDD が DRAM 内の要求された 512 バイトセクターを更新
- 4) HDD がメディアの 4KB セクターを上書き

## パフォーマンス

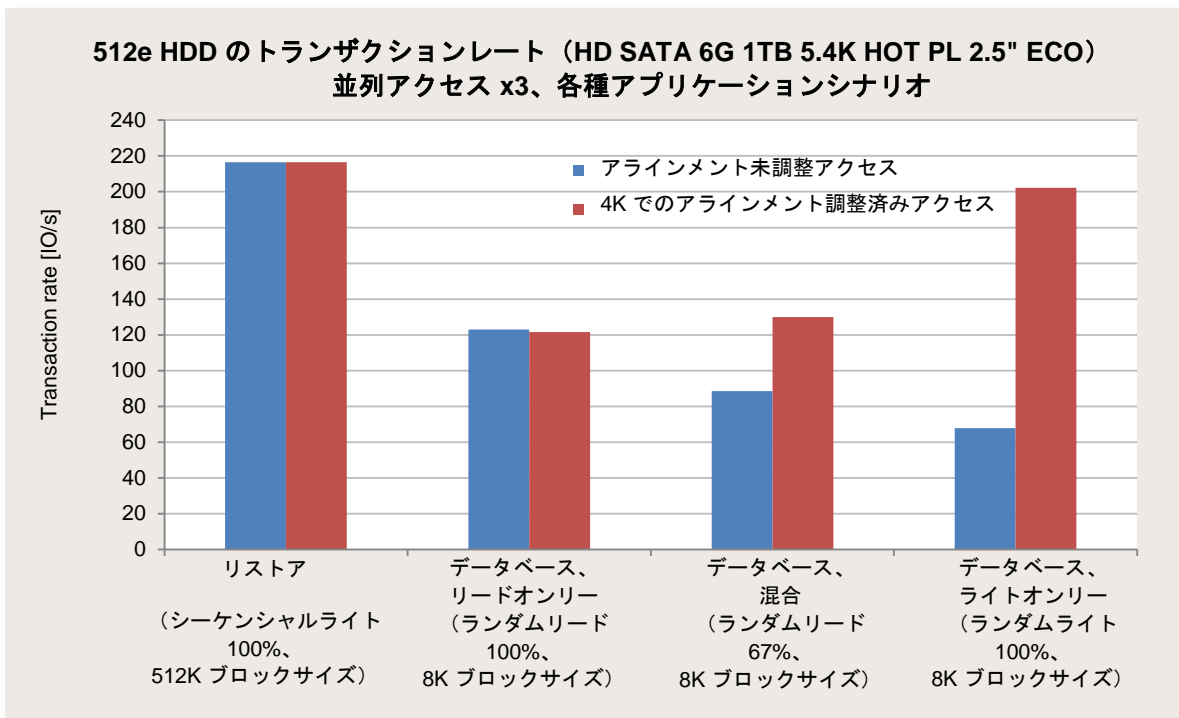
「読み取り-訂正-書き込み」のサイクルにより HDD 上で余分なヘッド動作が生じるため、パフォーマンスが低下します。このため、アプリケーションとオペレーティングシステム側で最高のパフォーマンスを実現させるには、物理 4K セクターの一部のみの書き込みアクセスを防止する必要があります。このようなケースには 2 つの場合があります。1 つは、ブロックサイズが 4K の完全な倍数にならない書き込みアクセスです。もう 1 つは、書き込みアクセスのブロックサイズが 4K の完全な倍数になりながら、物理セクターの境界に整列していない場合です。第 1 のケースは、今日のアプリケーションとオペレーティングでは書き込みブロックのサイズはほぼ常に 4K の倍数となることから、ほとんど問題になりません。このため、最適化の可能性のあるのは第 2 のケース、つまり書き込みアクセスのアラインメントの調整となります。

HDD の観点から書き込みアクセスのアラインメントを確保するには、特に以下の 2 つの要件を満たす必要があります。

- その 1 つは、書き込みが行われるパーティション (RAID アレイの場合はすべての HDD 上のパーツ) を物理 4k セクターと整合させることです。この前提条件は、例えば Windows Server 2008 以降に適用されます。このパーティションアラインメントは、パーティションの作成時から 4k に対応しているその後のアクセスでも、HDD の観点から 4k 整合アクセスを確保するための基本的な前提条件になります。
- もう 1 つは、パーティション作成時からすべての書き込みアクセスを 4k と整合させることです。これはより包括的な要件で、関連するすべてのソフトウェアコンポーネントに適用されます。つまり、書き込み操作の基本として使用できるように、パーティションの物理セクターサイズを正しく設定して伝える必要があります。このことは、オペレーティングシステムが開始するパーティションへの書き込みアクセス (およびこの方法で書き込みを行うアプリケーションからの「バッファード書き込み」) と「アンバッファード書き込み」の両方に適用されます。

物理 4k セクターを持つ HDD をさまざまな角度から考察した説明については、Microsoft の記事「[Advanced format \(4K\) disk compatibility update](#) (アドバンスドフォーマット (4k) ディスクの互換性に関する最新情報)」を参照してください。この記事では、物理 4k セクターに対するソフトウェアの適合性についても説明しています。

次の図は、アラインメント調整済アクセスとアラインメント未調整アクセスの場合の 512e HDD のパフォーマンスの違いを示しています。



100% 読み取りの場合は、シーケンシャルアクセスとランダムアクセスともパフォーマンスに違いはありません。アラインメント調整済みアクセスが有利になるのは、書き込み時のランダムアクセスのみとなっています。図の 3 列目のグループの例（データベースに典型的な頻繁なアクセスプロファイルに相当）では、書き込みが 33% を占めています。このため、アラインメント調整済みアクセスでは、書き込み率が大きいほどその利点も大きくなります。100% ライトのランダム負荷プロファイルでは、整合アクセスによって不整合アクセスの実に 3 倍のパフォーマンス優位性が得られます。これは図の 4 列目のグループが示しています。

補足 :

一部のハードディスクでは、追加の内部メカニズムによってアラインメント未調整書き込みアクセスの影響を緩和しています (例、Seagate 社の「[SmartAlign](#)」テクノロジー)。

## 高パフォーマンスサポートを実現するための 512e ハードディスクの推奨構成

最適なパフォーマンスを得るため、物理セクターサイズに関する情報を確認して、すべてのコンポーネントに正しく適用する必要があります。これは、ハードディスクコントローラー、ドライバおよびオペレーティングシステムが次の 3 つの表の条件を満たす場合に、通常オペレーティングシステムとアプリケーションによって実行される「バッファ書き込み」の場合に保証されます。

ディスクコントローラー			
タイプ	発注コード	説明	ファームウェアパッケージバージョン
RAID、PCIe スロット	S26361-F3554-E512/L512 S26361-F4481-E1/L1 S26361-F3593-E201/L201	RAID Ctrl SAS 6G 5/6 512 MB (D2616) PY SAS RAID Mezz Card 6Gb RAID Ctrl SAS 6G 8Port ex 512MB LP LSI	≥ 12.12.0~0129
	S26361-F3554-E8/L8	RAID Ctrl SAS 6G 0/1 (D2607)	≥ 20.10.1~0120
	S26361-F3669-E1/L1 S26361-F3669-E3/L3 S26361-F4531-E512/L512 S26361-F4531-E513/L513 S26361-F4531-E100/L100 S26361-F4531-E300/L300 S26361-F3713-E201/L201 S26361-F3713-E203/L203	RAID Ctrl SAS 6G 5/6 1GB (D3116) RAID Ctrl SAS 6G 5/6 1GB (D3116C) PY SAS RAID HDD Module PY SAS RAID HDD Module 3.0 PY SAS RAID HDD Module w/o cache PY SAS RAID HDD Module w/o cache 3.0 RAID Ctrl SAS 6G 8Port ex 1GB LP LSI RAID Ctrl SAS 6G 8Port ex 1GB LP LSI V3	現在のコントローラは 512e HDD に最適化されていません
RAID、オンボード	(ベースユニットに含まれる)	オンボードコントローラー、RAID モード	ファームウェアによる影響なし
非 RAID	S26361-F4480-E1 S26361-F3554-E118	PY SAS HBA Mezz Card 6Gb PSAS CP200i	ファームウェアによる影響なし

オペレーティングシステム	ドライババージョン
Windows タイプ	PCIe スロットコントローラー (RAID モード) : Megasas ≥ 6.505.5 オンボードコントローラー (RAID モード) : MegaSR ≥ 15.02.2013.0425 コントローラー (非 RAID モード) : すべて
Linux タイプ	PCIe スロットコントローラー (RAID モード) およびコントローラー (非 RAID モード) : ネイティブサポート オンボードコントローラー (RAID モード) : MegaSR ≥ 15.02.2013.0425

オペレーティングシステム	
名前	追加の要件
Windows Server 2008	<a href="#">KB2553708</a>
Windows Server 2008 R2	<a href="#">KB982018</a> または <a href="#">Service Pack 1</a>
Windows Server 2012	-
RHEL	≥ バージョン 5
SLES	≥ バージョン 11 SP2

国または販売地域によっては、一部のコンポーネントが利用できない場合があります。

## コメント :

- 表に記載されていない構成はすべて 512e HDD を機能的にサポートしますが、最適なパフォーマンスが得られません。
- リリースされているコントローラーとハードディスクの組み合わせについては、システム構成をご覧ください。
- オペレーティングシステムのインストール時に最適なパフォーマンスを確保するには、上記のソフトウェアコンポーネントをインストール手順の適切な段階で利用できる状態にする必要があります。以下のセクションで、PRIMERGY サーバ用の標準的な Windows バージョンなどでこの手順を実行する方法を説明します。
- 通常のオペレーティングシステム環境以外の稼働条件（例、イメージングツールを使用したパーティション作成またはハードディスク全体のクローン作成など）では、最適なパフォーマンスを実現するには、これらツールも 512e をサポートする必要があることに注意してください。
- すべてのアプリケーションにおいて、「バッファなし書き込み」での動作で、512e HDD へのアクセスで最適な書き込みパフォーマンスを得るには、物理セクターサイズを考慮する必要があります。



## Windows オペレーティングシステムのインストール

上記のホットフィックスとドライバは、通常、オペレーティングシステムのインストール後だけにインストールします。オペレーティングシステムを 512e-HDD にインストールする際の所要時間を最小限に抑えるには、インストール手順のできるだけ早い段階で、関連する追加ソフトウェアコンポーネントを利用できる状態にする必要があります。以下の表は、この「できるだけ早い段階で利用可能にする」という要件を満たすために、PRIMERGY サーバ用の標準的な Windows オペレーティングシステムバージョンで利用できるオプションの例をまとめたものです。

この表では、「ServerView Installation Manager」(SVIM) も考慮しています。この PRIMERGY サーバに付属のソフトウェアは、インストール手順の適切な段階で利用できる、追加のホットフィックスとドライバ (512e-HDD サポート用も含む) も提供します。以下で使われている「インストールソース」という用語は、オペレーティングシステムのインストール媒体をローカルまたはネットワーク経由で提供するさまざまなオプションを意味する一般的な用語です。

オペレーティングシステムのバージョン	使用するソフトウェアコンポーネントとハイパフォーマンスインストール (および最適なコントローラードライバ) に関する情報	
	OS 用	コントローラードライバ用
Windows Server 2008	SVIM バージョン 11.13.08 以降とすべてのインストールソース  または すべてのインストール段階に対応するホットフィックス <a href="#">KB2553708</a> が組み込まれたインストールソース *)	SVIM バージョン 11.13.08 以降とすべてのインストールソース  または 上記の <a href="#">ドライババージョン</a> 表に準じた推奨コントローラードライバ。このドライバは、インストール中に「Load Driver (ドライバの読み込み)」を選択して (または F6 を押して)、「Where do you want to install? (インストール場所を指定してください)」というメッセージが表示されたときに追加します。
Windows Server 2008 R2	SVIM バージョン 11.13.08 以降とすべてのインストールソース  または すべてのインストール段階に対応するホットフィックス <a href="#">KB982018</a> が組み込まれたインストールソース *)  または <a href="#">Service Pack 1</a> が組み込まれたインストールソース *)	
Windows Server 2012	-	
Windows Server 2012 R2	-	

\*) 適切な管理ツール (「Windows 自動インストールキット」(Windows AIK または WAIK) から利用できる「Deployment Image Servicing and Management」(DISM)、または「Windows アセスメント & デプロイメントキット」(Windows ADK)) を使用することで、Windows オペレーティングシステムバージョンのインストールイメージを元のイメージから拡張し、選択したパッケージを取り込むことができます。

以下の例は、オペレーティングシステムを 512e-HDD にインストールする際にどれだけ時間を短縮できるかを示すものです。ここでは、オンボードコントローラに接続した「HD SATA 6G 500GB 7.2K HOT PL 3.5" ECO」ハードディスクをサンプルとして使用しています。

ソフトウェアコンポーネント	インストールの所要時間
Windows Server 2008 R2、 ホットフィックス <a href="#">KB982018</a> が組み込まれていないインストールソース、 コントローラードライバ < 15.02.2013.0425	約 110 分
Windows Server 2008 R2、 ホットフィックス <a href="#">KB982018</a> が組み込まれているインストールソース、 コントローラードライバ ≥ 15.02.2013.0425	約 40 分

つまり、インストール時間を最小限に抑えるには、オペレーティングシステムとコントローラードライバの両方の前提条件を満たす必要があります。



## 関連資料

### PRIMERGY システム

<http://primergy.com/>

### ディスク I/O パフォーマンスの基本

<http://docs.ts.fujitsu.com/dl.aspx?id=35801735-a223-491a-a879-43f506444366>

### HDD 用「アドバンスフォーマット」テクノロジー

Seagate テクノロジーペーパー :

アドバンスフォーマットハードドライブ向け SmartAlign™ テクノロジー

[http://www.seagate.com/docs/pdf/whitepaper/tp615\\_smartalign\\_for\\_af\\_4k.pdf](http://www.seagate.com/docs/pdf/whitepaper/tp615_smartalign_for_af_4k.pdf)

### 「アドバンスフォーマット」HDD を使用したオペレーティングシステムとアプリケーションの互換性

Advanced format (4K) disk compatibility update

[http://msdn.microsoft.com/en-us/library/windows/desktop/hh848035\(v=vs.85\).aspx](http://msdn.microsoft.com/en-us/library/windows/desktop/hh848035(v=vs.85).aspx)

### Windows Server オペレーティングシステム用のホットフィックスとサービスパック

Hotfix KB2553708 for Windows Server 2008

<http://support.microsoft.com/kb/2553708/en-us>

Hotfix KB982018 for Windows Server 2008 R2

<http://support.microsoft.com/kb/982018/en-us>

Service Pack 1 for Windows Server 2008 R2

<http://support.microsoft.com/kb/976932/en-us>

### PRIMERGY のパフォーマンス

<http://www.fujitsu.com/fts/products/computing/servers/primergy/benchmarks/>

### PC サーバ PRIMERGY (プライマジー)

<http://jp.fujitsu.com/platform/server/primergy/>

## お問い合わせ先

### 富士通

Web サイト : <http://jp.fujitsu.com/>

### PRIMERGY のパフォーマンスとベンチマーク

<mailto:primergy.benchmark@ts.fujitsu.com>

© Copyright 2014 Fujitsu Technology Solutions. Fujitsu と Fujitsu ロゴは、富士通株式会社の日本およびその他の国における登録商標または商標です。その他の会社名、製品名、サービス名は、それぞれ各社の登録商標または商標です。知的所有権を含むすべての権利は弊社に帰属します。製品データは変更される場合があります。納品までの時間は在庫状況によって異なります。データおよび図の完全性、事実性、または正確性について、弊社は一切の責任を負いません。本書に記載されているハードウェアおよびソフトウェアの名称は、それぞれのメーカーの商標等である場合があります。第三者が各自の目的でこれらを使用した場合、当該所有者の権利を侵害することがあります。

詳細については、<http://www.fujitsu.com/fts/resources/navigation/terms-of-use.html> を参照してください。

2014-01-29 WW JA